LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-PROC-490083

Smoothed Aggregation Spectral Element Agglomeration AMG: SA-pAMGe

M. Brezina, P. S. Vassilevski

June 29, 2011

8th International Conference on "Large-Scale Scientific Computations"
Sozopol, Bulgaria
June 6, 2011 through June 10, 2011

Disclaimer

Smoothed aggregation spectral element agglomeration AMG: SA-ρAMGe ⋆ ⋆⋆

Marian Brezina[1] and Panayot S. Vassilevski[2]

[1] Department of Applied Mathematics, Campus Box 526, University of Colorado at Boulder, Boulder, CO 80309-0526, U.S.A.,
marian.brezina@gmail.com

[2] Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, P.O. Box 808, L-560, Livermore, CA 94550, U.S.A.
panayot@llnl.gov

Abstract. A two–level smoothed aggregation (or SA) scheme with tentative coarse space constructed by spectral element agglomeration method is shown to provide weak–approximation property in a weighted L_2–norm. The resulting method utilizing efficient (e.g., polynomial) smoothers is shown to have convergence factor independent of both the coarse and fine–grid mesh–sizes, as well as, to be independent of the contrast (i.e., possible large jumps in the PDE coefficient) for second order elliptic problems discretized on general unstructured meshes. The method allows for multilevel extensions. Presented numerical experiments exhibit behavior in agreement with the developed theory.

1 Introduction

This paper deals with the construction of two-grid iterative methods (or preconditioners) for solving elliptic problems with high contrast coefficient resolved only by a fine-grid discretization. The aim is to prove two–grid convergence bounds independent both of the fine and coarse mesh size as well as independent of the contrast. There are recent two-level results, [GE10], see also [SVZ], that deal with the construction of coarse spaces when used in combination with overlapping Schwarz method achieving this goal. The construction of the coarse spaces exploit local (element-based) procedures to construct local (overlapping) coarse spaces. These local spaces are patched using partition of unity. Similar construction goes back to the element–based algebraic multigrid (or AMGe) in [ρAMGi]. The latter method was modified in [ρAMGii], see also [LV08], to avoid the use of partition of unity and to allow for more straightforward multilevel extension

⋆ This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

⋆⋆ This work of the first author was sponsored by the Department of Energy under grant numbers DE-FG02-03ER25574 and DE-FC02-06ER25784, Lawrence Livermore National Laboratory under contract numbers B568677, and the National Science Foundation under grant numbers DMS-0621199, DMS-0749317, and DMS-0811275.

by applying recursion. So far this is the only known genuine multilevel AMGe method that does not lead to potentially highly ill-conditioned coarse bases (an issue arising from the partition of unity formed over overlapping domains used in a multilevel setting).

The present paper can be viewed as an extension of an early report [BHMV] from the point of view of handling problems with high contrast. That report contains also element-based construction of local coarse-spaces which are non–overlapping (or in terms of finite element functions they have minimal $\mathcal{O}(h)$ overlap). To improve the energy stability of the coarse basis, the polynomial smoothing idea from the multigrid smoothed aggregation (or SA) method is applied. In the present paper, we review this approach and also propose an alternative construction of the local (non–overlapping) spaces. We present a different analysis than [BHMV] following [BVV] to address both the high contrast coefficient and the aggressive coarsening. The two–level algorithm allows for multilevel extensions along the lines of [LV08] which provides an alternative to the formal multilevel extensions of the two–level methods exploiting partition of unity in [GE10] and [SVZ].

The remainder of the present paper is structured as follows. In Section 2 we provide two general conditions for the smoother and properties of the associated coarse space and prove that they ensure bounded two–level convergence factor. Section 3 contains the main results of the present paper. There, we review and extend the main results from [BHMV] paying attention to the high contrast coefficient and aggressive coarsening. Finally, Section 4 contains some numerical illustration of two- and multilevel results for problems with high–contrast.

2 A two-level convergence result

Consider the product iteration method involving "pre-smoothing" based on solving with M, coarse–grid solution involving restriction based on P^T, coarse-grid correction based on solving with $A_c = P^T A P$ and interpolation based on P, and finally using "post–smoothing" based on solving with M^T, represented by the following error-propagation matrix

$$E_{TL} \equiv (I - M^{-T}A)(I - PA_c^{-1}P^T A)(I - M^{-1}A). \tag{1}$$

Solving the system of linear equations, $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b}$ is a given right-hand side and $\mathbf{x}_0$ an initial iterate, one iteration of the two-level process is described by the following algorithm:

Algorithm 1 (Two–level iteration algorithm)

(i) *"Pre–smooth:" compute* $\mathbf{y}$ *from* $M(\mathbf{y} - \mathbf{x}_0) = \mathbf{b} - A\mathbf{x}_0$.
(ii) *Restrict defect* $\mathbf{r}_c = P^T(\mathbf{b} - A\mathbf{y})$.
(iii) *Solve for a coarse–grid correction* $\mathbf{x}_c$ *the coarse-grid problem* ($A_c = P^T A P$),

$$A_c \mathbf{x}_c = \mathbf{r}_c.$$

(iii) Interpolate and update fine–grid iterate, $\mathbf{z} = \mathbf{y} + P\mathbf{x}_c$.
(iv) "Post–smooth", i.e., solve for $\mathbf{x}_{TL}$, $M^T(\mathbf{x}_{TL} - \mathbf{z}) = \mathbf{b} - A\mathbf{z}$.

The mapping $\mathbf{b} \mapsto \mathbf{x}_{TL} = B_{TL}^{-1}\mathbf{b}$ for $\mathbf{x}_0 = 0$ defines the inverse of the two–level operator B_{TL}. Alternatively, formula (1) defines an approximate inverse (or preconditioner) to A, B_{TL}^{-1}, from the identity $I - B_{TL}^{-1}A = E_{TL}$. If M is A–convergent, that is $\|I - A^{\frac{1}{2}}M^{-1}A^{\frac{1}{2}}\| < 1$ (or equivalently, $M + M^T - A$ is s.p.d.) we then have that B_{TL}^{-1} is s.p.d. and B_{TL} can be characterized by the following two-level version of the XZ-identity, cf. [Va08]:

$$\mathbf{v}^T B_{TL}\mathbf{v} = \min_{\mathbf{v}=\mathbf{v}_f+P\mathbf{v}_c} \left[\mathbf{v}_c^T A_c\mathbf{v}_c + \left(\mathbf{v}_f - M^{-T}AP\mathbf{v}_c\right)^T \overline{M}\left(\mathbf{v}_f - M^{-T}AP\mathbf{v}_c\right)\right].$$

Here $\overline{M} = M\left(M + M^T - A\right)^{-1}M^T$ is the symmetrized smoother. For any A-convergent smoother M and $A_c = P^TAP$, we have

$$\mathbf{v}^T A\mathbf{v} \le \mathbf{v}^T B_{TL}\mathbf{v}.$$

In order to establish convergence properties of the two-level process, our goal is to estimate the constant, K_{TL}, in the upper bound,

$$\mathbf{v}^T B_{TL}\mathbf{v} \le K_{TL}\mathbf{v}^T A\mathbf{v},$$

which is equivalent to $0 \le \mathbf{v}^T AE_{TL}\mathbf{v} \le \varrho_{TL}\ \|\mathbf{v}\|_A^2$, where $K_{TL} = \frac{1}{1-\varrho_{TL}}$. For this we use the above two–level (XZ-) identity by selecting a proper coarse–grid approximation $P\mathbf{v}_c$ to any given $\mathbf{v}$ which gives

$$\mathbf{v}^T B_{TL}\mathbf{v} \le \mathbf{v}_c^T A_c\mathbf{v}_c + 2\|\mathbf{v} - P\mathbf{v}_c\|_M^2 + 2\mathbf{v}_c^T P^T A\left(M + M^T - A\right)^{-1}AP\mathbf{v}_c. \quad (2)$$

Consider the particular case when A is a s.p.d. sparse matrix coming from a finite element discretization of an elliptic PDE posed on a domain $\Omega \subset \mathbf{R}^d$, (plane polygon for $d = 2$, or polytope in the case of $d = 3$). We use quasi-uniform mesh $\mathcal{T}_h$ with mesh size h and respective finite element space V_h. Also, let $\mathcal{N}_h$ be the set of Lagrangian degrees of freedom, that in the case of piecewise linear finite element space V_h coincide with the values at the vertices $\{\mathbf{x}_i\}$ of the elements in $\mathcal{T}_h$. The corresponding elliptic bilinear form reads

$$a(u,\ \varphi) = \int_\Omega k(\mathbf{x})\nabla u \cdot \nabla\varphi\ d\mathbf{x}, \quad (3)$$

where $k = k(\mathbf{x})$ is a given positive coefficient. The coefficient may admit very large jumps in certain parts of Ω which are assumed resolved by the fine-grid $\mathcal{T}_h$, i.e., $k(\mathbf{x})$ varies smoothly within each fine–grid element $\tau \in \mathcal{T}_h$. Assume also that we have created a coarse finite element space $V_H \subset V_h$ where H reflects the characteristic diameter of the support of the coarse-basis functions. We also assume that the coarse space satisfies the following approximation property: For any $v \in V_h$ there is a $v_H \in V_H$ such that uniformly in h and H, we have

$$H^{-2}\|v - v_H\|_0^2 + a(v - v_H,\ v - v_H) \le Ca(v,\ v). \quad (4)$$

Here, $\|.\|_0$ is the k–weighted L_2–norm, i.e., $\|\varphi\|_0^2 = \int\limits_{\Omega} k(\mathbf{x})\ \varphi^2(\mathbf{x})\ d\mathbf{x}$. Without loss of generality, we may assume that k is piecewise constant with respect to the elements of $\mathcal{T}_h$. In that case, we have the following norm equivalence for any finite element function $v \in V_h$ and its nodal coefficient vector $\mathbf{v} = (v(\mathbf{x}_i))_{\mathbf{x}_i \in \mathcal{N}_h}$

$$\|v\|_0^2 = \int\limits_{\Omega} k(\mathbf{x})\ v^2(\mathbf{x})\ d\mathbf{x} \simeq \sum_{\mathbf{x}_i \in \mathcal{N}_h} d_i\ v^2(\mathbf{x}_i) = \|\mathbf{v}\|_{D_G}^2. \tag{5}$$

Here, $D_G = \text{diag}\ (d_i)$ is the diagonal of the k–weighted mass matrix G corresponding to the finite element space V_h and the weighted L_2 bilinear form $(k\ \cdot,\ \cdot)$. We have the following simple inverse inequality

$$a(v,\ v) \leq C_I h^{-2}\ \|v\|_0^2,$$

with a constant C_I independent of the coefficient $k = k(\mathbf{x})$ (which is assumed piecewise constant with respect to the elements $\tau \in \mathcal{T}_h$). Now, consider D_A– the diagonal of A and D_G–the diagonal of the weighted mass matrix G. The following norm equivalence

$$\mathbf{v}^T D_A \mathbf{v} \simeq h^{-2}\ \mathbf{v}^T D_G \mathbf{v}, \tag{6}$$

is easily seen to be uniform with respect to both $h \mapsto 0$ and the contrast (jumps of k). This can be proved directly based on an element-by-element local comparison of the two quadratic forms. Here, we use the assumption that $\mathcal{T}_h$ is quasi-uniform.

We consider smoothers M that possess the following "smoothing" property:

$$\mathbf{v}^T \overline{M} \mathbf{v} \leq \beta \left[\mathbf{v}^T A \mathbf{v} + \frac{b}{(H/h)^2} \|\mathbf{v}\|_{\Lambda}^2 \right]. \tag{7}$$

Here Λ is a given s.p.d. matrix. b satisfies $\|\Lambda^{-\frac{1}{2}} A \Lambda^{-\frac{1}{2}}\| \leq b$. In our case Λ will be spectrally equivalent to D_A. Consider the polynomial smoother M

$$M^{-1} = \left[I - p_\nu(b^{-1}\Lambda^{-1}A) \right] A^{-1},$$

where $p_\nu(t)$, $p_\nu(0) = 1$, is a properly chosen polynomial of degree ν, (cf., [BVV]), or Section 4 for details). The following estimate holds ([BVV])

$$\mathbf{x}^T \overline{M} \mathbf{x} \leq \beta\ \left(\mathbf{x}^T A \mathbf{x} + \frac{b}{(2\nu+1)^2}\ \|\mathbf{x}\|_{\Lambda}^2 \right). \tag{8}$$

Thus, if we choose the polynomial degree, ν, so that $2\nu + 1 \simeq \frac{H}{h}$, we obtain estimate (7).

Note that when $\Lambda = D_A$ (the diagonal of A) (or spectrally equivalent to D_A) this would make $b \simeq \|D_A^{-\frac{1}{2}} A D_A^{-\frac{1}{2}}\| = \mathcal{O}(1)$ (see Remark 1 below).

For the same polynomial smoother M, it has been shown in [BVV] that the following coercivity bound holds:

$$\mathbf{v}^T (M + M^T - A) \mathbf{v} \geq \alpha\ \mathbf{v}^T A \mathbf{v}. \tag{9}$$

The constant α in (9) is independent of both h and H.

Now, we are ready to prove the following two-level convergence result.

Theorem 1. *Assume that the A–convergent smoother M and its respective symmetrized version $\overline{M}$ satisfy "smoothing property" (7) where D_A is the diagonal of A, and in addition let the coercivity estimate (9) hold. Under the assumption of norm–equivalence of D_A and the diagonal of the weighted mass matrix G, $\|\mathbf{v}\|_{D_A} \simeq h^{-1} \|\mathbf{v}\|_{D_G}$, assuming that the coarse finite element space V_H ensures the approximation property (4), which rewritten in a matrix-vector form (letting $P\mathbf{v}_c$ be the coefficient vector of $v_H \in V_H$) reads,*

$$H^{-2}\|\mathbf{v} - P\mathbf{v}_c\|_G^2 + \|\mathbf{v} - P\mathbf{v}_c\|_A^2 \leq C_a \ \|\mathbf{v}\|_A^2, \tag{10}$$

the following convergence estimate holds:

$$\frac{1}{1-\varrho_{TL}} = K_{TL} \leq \mathit{Const} \simeq 2\beta \ \max\{1,\ b\} C_a + (1+\frac{2}{\alpha})\ (1+\sqrt{C_a})^2.$$

That is, K_{TL} is bounded uniformly with respect to both h and H.

Proof. Using the approximation property (4) together with the assumed norm–equivalences shows

$$h^2\|\mathbf{v} - P\mathbf{v}_c\|_{D_A}^2 \simeq \|\mathbf{v} - P\mathbf{v}_c\|_{D_G}^2 \simeq \|\mathbf{v} - P\mathbf{v}_c\|_G^2 \leq C_a H^2 \ \|\mathbf{v}\|_A^2.$$

By the triangle inequality, we also have that the coarse-grid interpolant $P\mathbf{v}_c$ is bounded in energy, i.e.,

$$\|\mathbf{v}_c\|_{A_c}^2 = \|P\mathbf{v}_c\|_A^2 \leq (1+\sqrt{C_a})^2 \ \|\mathbf{v}\|_A^2.$$

Now, from (2), using the smoothing property of M, (7), and the assumed coercivity (9), we obtain

$$\begin{aligned}
\mathbf{v}^T B_{TL}\mathbf{v} &\leq \mathbf{v}_c^T A_c \mathbf{v}_c + 2\|\mathbf{v} - P\mathbf{v}_c\|_{\overline{M}}^2 + 2\mathbf{v}_c^T P^T A \left(M + M^T - A\right)^{-1} AP\mathbf{v}_c \\
&\leq \mathbf{v}_c^T A_c \mathbf{v}_c + \frac{2}{\alpha}\|P\mathbf{v}_c\|_A^2 + 2\beta \ \left(\|\mathbf{v} - P\mathbf{v}_c\|_A^2 + \frac{b}{(H/h)^2} \ \|\mathbf{v} - P\mathbf{v}_c\|_{D_A}^2\right) \\
&\simeq (1+\frac{2}{\alpha}) \ \|\mathbf{v}_c\|_{A_c}^2 + 2\beta \ \left(\|\mathbf{v} - P\mathbf{v}_c\|_A^2 + b\ H^{-2} \ \|\mathbf{v} - P\mathbf{v}_c\|_{D_G}^2\right) \\
&\simeq (1+\frac{2}{\alpha}) \ \|\mathbf{v}_c\|_{A_c}^2 + 2\beta \ \left(\|\mathbf{v} - P\mathbf{v}_c\|_A^2 + b\ H^{-2} \ \|\mathbf{v} - P\mathbf{v}_c\|_G^2\right) \\
&\leq (1+\frac{2}{\alpha}) \ \|\mathbf{v}_c\|_{A_c}^2 + 2\beta \ \max\{1,\ b\} \ C_a \ \|\mathbf{v}\|_A^2 \\
&\leq \left[2\beta \ \max\{1,\ b\} C_a + (1+\frac{2}{\alpha}) \ (1+\sqrt{C_a})^2\right] \ \mathbf{v}^T A\mathbf{v}.
\end{aligned}$$

Remark 1. We remark that for any s.p.d. sparse matrix A, we have the simple estimate

$$\mathbf{v}^T A\mathbf{v} \leq \kappa \ \mathbf{v}^T D_A \mathbf{v},$$

where $\kappa \geq 1$ stands for the maximal number of non-zero entries per row of A. This shows that $\|D_A^{-\frac{1}{2}} A D_A^{-\frac{1}{2}}\| \leq \kappa$ and we can let $b = \kappa$. Therefore, under our assumption of quasi-uniform mesh, we may assume that in the estimates derived in Theorem 1, b is a constant that is both mesh-independent and independent of the coefficient $k(\mathbf{x})$.

Remark 2. Theorem 1 shows that K_{TL} will be bounded independently of possible large jumps in the PDE coefficient $k = k(\mathbf{x})$ as long as the approximation property (10) holds with a constant C_a independent of those jumps.

3 A spectral, element agglomeration based, construction of coarse spaces with weak approximation property

In the present section, we review an approach of constructing coarse spaces with weak approximation property (4), or its matrix–vector form (10), that requires knowledge of local element matrices. For details, we refer to the early papers [BHMV], [ρAMGi], [ρAMGii], and to some more recent ones that explicitly address the issue of high-contrast coefficients, [GE10], [SVZ]. Most of these methods are inherently two–level with the exception of [ρAMGii]. The methods, for example, in the more recent papers [GE10], [SVZ] can in principle be applied recursively, however issues related to the use of partition of unity have to be resolved since the resulting coarse matrices may get fairly ill-conditioned due to the (nearly) linear dependence of the respectively constructed coarse bases.

We consider the model second order elliptic bilinear form (3) and a quasi-uniform fine–grid triangulation $\mathcal{T}_h$ that is fine-enough to resolve the possible large jumps of the coefficient $k = k(\mathbf{x})$. Without loss of generality, we may assume that $k = k(\mathbf{x})$ is piecewise constant with respect to the elements of $\mathcal{T}_h$. We now present two spectral aggregation/element agglomeration based approaches to construct accurate enough coarse spaces. We assume that a set $\mathcal{T}_H$ of non–overlapping agglomerated elements $\{T\}$ has been constructed. This means that each T is a (connected) union of fine–grid elements $\{\tau\}$. Let the characteristic mesh size (diameter of $T \in \mathcal{T}_H$) be of order H. We do not assume that H is comparable to the fine–grid mesh–size h. For each T, we assemble the local stiffness matrix A_T and the local weighted mass matrix G_T. In addition to the set $\mathcal{T}_H$ of agglomerated elements T, we assume that we have respective aggregates $\{\mathcal{A}_i\}_{i=1}^{n_A}$ where each $\mathcal{A} = \mathcal{A}_i$ is contained in a unique T. The set $\{\mathcal{A}_i\}$ provides a non–overlapping partition of the set $\mathcal{N}_h$ of fine–degrees of freedom. The aggregates are easily constructed based on $\mathcal{T}_H$, by assigning the interface nodes, shared by multiple $\bar{T}$, $T \in \mathcal{T}_H$, to only one of the aggregates. We note that is is advantageous to assign the interface degrees of freedom to the aggregate with the largest value of the coefficient, k.

3.1 The construction of tentative prolongator

In [BHMV] the following construction was used. Solve the generalized eigenvalue problem (we use the equivalent eigenvalue problem as proposed in [GE10])

$$A_T \mathbf{q}_k = \overline{\lambda}_k G_T \mathbf{q}_k, \; k = 1, \; \ldots, \; n_T, \tag{11}$$

where n_T is the number of fine–degrees of freedom in T. By choosing the first m_T eigenvectors in the lower part of the spectrum of $G_T^{-1} A_T$, we form the

rectangular matrix $Q_T = [\mathbf{q}_1, \ldots, \mathbf{q}_{m_T}]$. Then, we extract the rows of Q_T with row–indices from the aggregate $\mathcal{A}$ (where $\mathcal{A} \subset T$) and form $Q_{\mathcal{A}}$. Finally, using for example SVD, we can form a linearly independent set from the columns of $Q_{\mathcal{A}}$. The resulting matrix $\widehat{P}_{\mathcal{A}}$ has orthogonal (hence linearly independent) columns. The global tentative prolongator is simply the block-diagonal matrix, with the blocking corresponding to individual aggregates,

$$\widehat{P} = \begin{bmatrix} \widehat{P}_{\mathcal{A}_1} & 0 & \ldots & 0 \\ 0 & \widehat{P}_{\mathcal{A}_2} & \ldots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \ldots & 0 & \widehat{P}_{\mathcal{A}_{n_A}} \end{bmatrix}. \tag{12}$$

Since for quasi-uniform mesh $\mathcal{T}_h$ it is easily seen that $h^2 D_T$, the scaled diagonal of A_T, and G_T, are uniformly spectrally equivalent, in practice we can instead solve (as in [BHMV]) the generalized eigenvalue problem equivalent to (11),

$$A_T \mathbf{q}_k = \lambda_k D_T \mathbf{q}_k, \; k = 1, \ldots, n_T. \tag{13}$$

Note that the respective eigenvalues are related as

$$\overline{\lambda}_k \simeq h^{-2} \lambda_k. \tag{14}$$

It is straightforward to see the following local estimate

$$\|\mathbf{v} - \widehat{P}\mathbf{v}_c\|^2_{D_T} \leq \frac{1}{\lambda_{m_T+1}} \|\mathbf{v}\|^2_{A_T}. \tag{15}$$

By summing up the local estimates, using the equivalence between $h^2 D_T$ and G_T, we arrive at the weak approximation property (essentially proven in [BHMV])

$$H^{-2} \|\mathbf{v} - \widehat{P}\mathbf{v}_c\|^2_G \leq \sigma \max_T \frac{h^2}{H^2 \lambda_{m_T+1}} \mathbf{v}^T A \mathbf{v}. \tag{16}$$

Here, σ is a constant independent of H, h and $k = k(\mathbf{x})$. Noticing now that $\frac{1}{\lambda_{m_T+1}}$ scales with $\left(\frac{H}{h}\right)^2$ (easily seen from (11), since $\overline{\lambda}_k$ scales as H^{-2} and $\lambda_k \simeq h^2 \overline{\lambda}_k$, see (14)), it follows that by choosing m_T appropriately (sufficiently large) a uniform bound with respect to both, H/h and the contrast (the size of jumps of $k = k(\mathbf{x})$ within each T) can be ensured. The size of m_T depends only on the number of subdomains of each T where k admits these large jumps. More detailed investigation on the optimal choice of m_T can be found in [GE10], see also [SVZ]. At any rate, an upper bound for m_T is the number of subdomains of T where k has substantially different values from the other subdomains.

3.2 Alternative construction of tentative prolongator

With the purpose to allow for multilevel extension, we propose the following alternative procedure. Based on A_T, we compute its reduced Schur complement

form $S_{\mathcal{A}}$ corresponding to the degrees of freedom of the aggregate $\mathcal{A}$, $\mathcal{A} \subset T$. Then, we solve the generalized eigenvalue problem

$$S_{\mathcal{A}}\mathbf{p}_k = \lambda_k D_{\mathcal{A}}\mathbf{p}_k, k = 1, \ \ldots, \ n_{\mathcal{A}},$$

where $D_{\mathcal{A}}$ is, for example, the diagonal of A_T restricted to $\mathcal{A}$. Here, $n_{\mathcal{A}}$ stands for the size of $\mathcal{A}$ (the number of degrees of freedom in $\mathcal{A}$). The tentative prolongator is simply the block–diagonal matrix $\widehat{P}$ with block-entries $\widehat{P}_{\mathcal{A}} = [\mathbf{p}_1, \ \ldots, \ \mathbf{p}_{m_{\mathcal{A}}}]$ for a sufficiently large $m_{\mathcal{A}} \leq n_{\mathcal{A}}$. In practice, we choose $m_{\mathcal{A}}$ such that for a given tolerance $\theta \in (0, 1]$,

$$\lambda_k \geq \theta \ \lambda_{\max} \text{ if } k > m_{\mathcal{A}} \ \ (\text{assuming } \lambda_s \leq \lambda_{s+1}). \tag{17}$$

It is straightforward to see the local estimates

$$\|\mathbf{v}_{\mathcal{A}} - \widehat{P}_{\mathcal{A}}\mathbf{v}^c_{\mathcal{A}}\|^2_{D_{\mathcal{A}}} \leq \left(\frac{1}{\lambda_{m_{\mathcal{A}}+1}}\right) \ \mathbf{v}^T_{\mathcal{A}} S_{\mathcal{A}}\mathbf{v}_{\mathcal{A}} \leq \left(\frac{1}{\lambda_{m_{\mathcal{A}}+1}}\right) \ \mathbf{v}^T_T A_T \mathbf{v}_T,$$

where $\mathbf{v}_T$ is any extension of $\mathbf{v}_{\mathcal{A}}$ (defined on $\mathcal{A}$) to a vector $\mathbf{v}_T = \begin{bmatrix} \mathbf{v}_{\mathcal{A}} \end{bmatrix}$ defined on T. We use here the fact that $S_{\mathcal{A}}$ is a Schur complement of the symmetric positive semi–definite matrix A_T. By summing up the local estimates over $\mathcal{A}$, the following global one is obtained (D_A is the diagonal of A)

$$\|\mathbf{v} - \widehat{P}\mathbf{v}_c\|^2_{D_A} \leq \left(\max_{\mathcal{A}} \ \frac{1}{\lambda_{m_{\mathcal{A}}+1}}\right) \ \mathbf{v}^T A \mathbf{v}. \tag{18}$$

Finally, using the uniform spectral equivalence between $h^2 D_A$ and D_G (and G), we obtain the desired weak approximation estimate of the form (16).

Above, D_A can be replaced with any other spectrally equivalent matrix D. For example, in Section 4, we consider a particular choice of D that makes $b = \|D^{-\frac{1}{2}} A D^{-\frac{1}{2}}\| = 1$.

3.3 The smoothed prolongator and its analysis

Based on the tentative prolongator $\widehat{P}$ and the matrix polynomial $S = s_\nu(b^{-1}D^{-1}A)$, where $b: \ \|D^{-\frac{1}{2}} A D^{-\frac{1}{2}}\| \leq b = \mathcal{O}(1)$ and

$$s_\nu(t) = (-1)^\nu \frac{1}{2\nu+1} \frac{T_{2\nu+1}(\sqrt{t})}{\sqrt{t}}, \tag{19}$$

with $T_l(t)$ denoting the Chebyshev polynomial of the first kind and degree l over the interval $[-1, +1]$, we define the actual smoothed aggregation prolongation matrix P as

$$P = S\widehat{P}.$$

The effect of prolongation smoothing is that it makes the prolongation operator, P, stable in the energy norm. More specifically, letting Q be the D-orthogonal projection onto the space Range($\widehat{P}$), we get

$$\|SQ\mathbf{v}\|_A \leq \|S\mathbf{v}\|_A + \|S(\mathbf{v} - Q\mathbf{v})\|_A \leq \|\mathbf{v}\|_A + \frac{b^{\frac{1}{2}}}{2\nu+1} \ \|\mathbf{v} - Q\mathbf{v}\|_D.$$

Here, we used the fact that $s_\nu^2(t) \in [0,1]$ for $t \in (0,1]$, and also its main property (cf., e.g., [Va08])

$$\sup_{t\in(0,\,1]} |\sqrt{t} s_\nu(t)| \le \frac{1}{2\nu+1}.$$

Using now the proven weak approximation property of $Q\mathbf{v}$, i.e., estimate (18), we arrive at the following final energy norm bound of SQ,

$$\|SQ\mathbf{v}\|_A \le \left[1 + \frac{b^{\frac{1}{2}}}{2\nu+1}\ \frac{H}{h}\ \left(\max_{\mathcal{A}}\ \frac{h^2}{H^2\lambda_{m_{\mathcal{A}}+1}}\right)^{\frac{1}{2}}\right]\ \|\mathbf{v}\|_A. \tag{20}$$

Choosing $\nu \simeq H/h$ ensures SQ being uniformly bounded in energy. Similarly,

$$\begin{aligned}\|\mathbf{v} - SQ\mathbf{v}\|_D &\le \|S(\mathbf{v}-Q\mathbf{v})\|_D + \|(I-S)\mathbf{v}\|_D \le \|(I-Q)\mathbf{v}\|_D + \|(I-S)\mathbf{v}\|_D \\ &= \|(I-Q)\mathbf{v}\|_D + \|D^{\frac{1}{2}}\left(I - s_\nu(b^{-1}D^{-1}A)\right)A^{-\frac{1}{2}}\|\|\mathbf{v}\|_A.\end{aligned}$$

The last matrix norm is estimated as follows

$$\|D^{\frac{1}{2}}\left(I - s_\nu(b^{-1}D^{-1}A)\right)A^{-\frac{1}{2}}\| = \|X^{-\frac{1}{2}}\left(I - s_\nu(b^{-1}X)\right)\|, \quad X = A^{\frac{1}{2}}D^{-1}A^{\frac{1}{2}}.$$

Therefore,

$$\begin{aligned}\|\mathbf{v} - SQ\mathbf{v}\|_D &\le \|(I-Q)\mathbf{v}\|_D + b^{-\frac{1}{2}} \sup_{t\in(0,\,1]} \frac{1-s_\nu(t)}{\sqrt{t}}\ \|\mathbf{v}\|_A \\ &\le \left(\max_{\mathcal{A}}\ \frac{1}{\sqrt{\lambda_{m_{\mathcal{A}}+1}}}\right)\ \|\mathbf{v}\|_A + \frac{1}{\sqrt{b}}\ C_\nu\ \|\mathbf{v}\|_A.\end{aligned}$$

Thus, using the equivalence $h^2 D \simeq G$, we have the final estimate

$$\begin{aligned}H^{-1}\|\mathbf{v} - SQ\mathbf{v}\|_G &\simeq \tfrac{h}{H}\ \|\mathbf{v} - SQ\mathbf{v}\|_D \\ &\le \tfrac{h}{H}\ \left(\left(\max_{\mathcal{A}}\ \frac{1}{\sqrt{\lambda_{m_{\mathcal{A}}+1}}}\right)\ \|\mathbf{v}\|_A + \frac{1}{\sqrt{b}}\ C_\nu\ \|\mathbf{v}\|_A\right) \\ &\le \left(\left(\max_{\mathcal{A}}\ \frac{h}{H\sqrt{\lambda_{m_{\mathcal{A}}+1}}}\right) + b^{-\frac{1}{2}}\ \frac{C_\nu}{2\nu+1}\ \frac{2\nu+1}{\frac{H}{h}}\right)\|\mathbf{v}\|_A.\end{aligned}$$

Since $\frac{C_\nu}{2\nu+1} = \mathcal{O}(1)$ (see [BVV]), by choosing $\nu \simeq \frac{H}{h}$, we obtain the desired uniform boundedness of $H^{-1}\|\mathbf{v} - SQ\mathbf{v}\|_G$ in terms of $\|\mathbf{v}\|_A$. We summarize:

Theorem 2. *The two–level spectral element agglomeration construction of tentative prolongator $\widehat{P}$ combined with the smoothed aggregation construction of the actual interpolation matrix $P = S\widehat{P}$ where $S = s_\nu(b^{-1}D^{-1}A)$ and s_ν is given in (19) for $\nu \simeq \frac{H}{h}$, ensures the weak approximation property with constant $\beta_1 = O(1)$:*

$$H^{-1}\|\mathbf{v} - SQ\mathbf{v}\|_G \le \beta_1\ \left(1 + \max_{\mathcal{A}}\ \left(\frac{h^2}{H^2\lambda_{m_{\mathcal{A}}+1}}\right)^{\frac{1}{2}}\right)\ \|\mathbf{v}\|_A.$$

Also, the energy stability property holds with constant $\beta_2 = O(1)$:

$$\|SQ\mathbf{v}\|_A \le \left(1 + \beta_2\ \max_{\mathcal{A}}\ \left(\frac{h^2}{H^2\lambda_{m_{\mathcal{A}}+1}}\right)^{\frac{1}{2}}\right)\ \|\mathbf{v}\|_A.$$

Based on Theorems 1–2, Remark 2 implies the following corollary.

Corollary 1. *The two–level spectral element agglomeration construction of tentative prolongator $\widehat{P}$ combined with the smoothed aggregation construction of the actual interpolation matrix as described in Theorem 2, exhibits convergence bound K_{TL} that is also contrast independent (in addition to the independence of both h and H).*

4 Numerical experiments

We have implemented a multilevel version of the proposed SA spectral AMGe algorithm to test the performance of both the two-level (TL) method and its multilevel (ML) extension. We use agglomeration algorithm that exploits the fine–grid vertex coordinates, as described in [BVV]. In this way, the aggregates after the first coarse level have complexity similar to a geometric multigrid.

At coarse levels we use as multigrid relaxation a simple block Gauss–Seidel smoothing with blocks corresponding to the respective aggregates. At the finest level, we use as multigrid relaxation the polynomial smoother from [VBT],

$$p_\nu(t) = (1 - T^2_{2\nu+1}(\sqrt{t}))s_\nu(t).$$

Its smoothing and coercivity properties, (8)-(9), have been analyzed in [BVV].

To select the local eigenvectors (needed in the construction of the tentative prolongator, see (17)), we used fairly small tolerance $\theta = 0.01$. All experiments were run as a stationary iteration method with stopping criteria that halts the iteration once the preconditioned residual norm was reduced by a factor of 10^{-6} relative to the initial one.

In the experiments, we have chosen the weighted ℓ_1–smoother $\Lambda = \text{diag}\,(d_i)_{i=1}^n$ where for the given s.p.d. matrix $A = (a_{ij})_{i,j=1}^n$ and any given positive weights $\{w_i\}$, we let $d_i = \sum\limits_j |a_{ij}| \frac{w_i}{w_j}$. For any given positive weights $\{w_i\}$, we have $\mathbf{v}^T A\mathbf{v} \leq \mathbf{v}^T \Lambda \mathbf{v}$. In the experiments, we have chosen $w_l = \sqrt{a_{ll}}$. In that case, we also have that Λ is spectrally equivalent to D_A (the diagonal of A), i.e., $\mathbf{v}^T D_A \mathbf{v} \leq \mathbf{v}^T \Lambda \mathbf{v} \leq \kappa\ \mathbf{v}^T D_A \mathbf{v}$, where κ is the maximum number of nonzero entries per row of A. We note that Theorem 1 applies with the above choice of Λ, since in its proof, we can replace D_A with any spectrally equivalent s.p.d. matrix Λ.

Our fine–grid problem is posed on the unit square domain Ω where the coefficient $k = k(\mathbf{x})$ takes two values: 1 and 10^c for various values of $c = -12, -9, -6, -3, 0, 1, 3, 6, 9,\ 12$. The region where $k = 10^c$ is shown in Figure 1 (right). The coefficient is resolved only on the finest mesh (see the middle of Figure 1) by using adaptive local refinement starting from an unstructured coarse mesh, shown in Figure 1 (left). The initial mesh does not resolve the coefficient. The distribution of the values 1 and 10^c is illustrated in Figure 1 (right).

Form Table 1, we can see that the two-level (TL) SA spectral AMGe exhibits convergence factors that are fairly insensitive with respect to the contrast for a

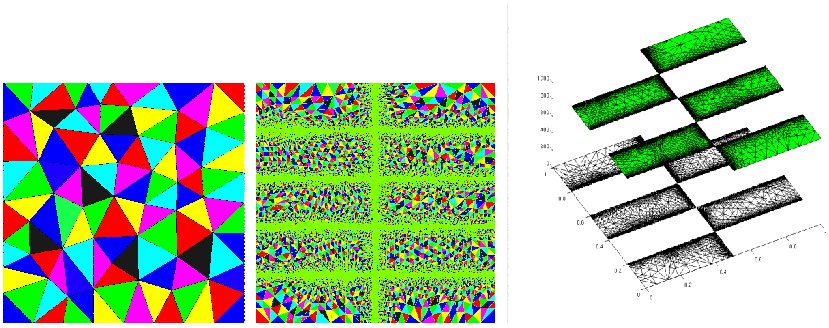

Fig. 1. Initial unstructured mesh (left) that does not resolve the discontinuous coefficient, its refinement (middle) that resolves the high-contrast discontinuous coefficient, and the actual distribution of the coefficient values 1 and 10^c, $c = 3$ (right).

c	# coarse grid dofs	n_{it}	ϱ
-12	2,729	53	0.74
-9	2,730	53	0.75
-6	2,730	51	0.74
-3	2,720	26	0.60
0	2,660	22	0.60
3	2,728	24	0.60
6	2,744	36	0.65
9	2,744	46	0.70
12	2,744	46	0.70

Table 1. Convergence factor (ϱ) and number of iterations (n_{it}) for the two–level SA spectral AMGe with $\nu = 6$ (degree of polynomial smoother $3\nu + 1 = 19$). The fine (triangular) mesh is fixed with $465,712$ elements and $233,499$ dofs. The coarse mesh consists of $1,330$ elements. The jump in the PDE coefficient is 10^c.

c	# coarse grid dofs	n_{it}^{TL}	ϱ_{TL}	n_{it}^{ML}	ϱ_{ML}
-12	11,191	57	0.73	77	0.79
-9	11,191	56	0.73	77	0.79
-6	11,191	46	0.70	67	0.78
-3	11,193	21	0.53	39	0.71
0	11,073	19	0.54	33	0.72
3	11,199	21	0.53	43	0.73
6	11,199	36	0.64	44	0.71
9	11,199	39	0.65	49	0.69
12	11,199	39	0.65	58	0.71

Table 2. Convergence factor (ϱ) and number of iterations (n_{it}) for both two-level (TL) and the multilevel (ML) SA spectral AMGe with $\nu = 6$ (degree of polynomial smoother $3\nu + 1 = 19$) at the finest level. At coarser level (in the ML case) block Gauss–Seidel smoother is used with blocks corresponding to the respective aggregates. The fine (triangular) mesh is fixed with $465,712$ elements and $233,499$ dofs. The coarse mesh consists of $5,250$ elements. The jump in the PDE coefficient is 10^c.

# fine grid dofs	# fine grid elements	# dofs at 1st coarse level	# elements at 1st coarse level	n_{it}^{ML}	ϱ_{ML}	n_{it}^{TL}	ϱ_{TL}	operator complexity
40,377	80,192	1,237	570	17	0.36	13	0.26	1.13
100,613	200,416	3,410	1,596	29	0.55	19	0.49	1.14
233,499	465,712	11,199	5,250	58	0.71	39	0.65	1.19
578,017	1,153,792	43,026	20,860	84	0.80	38	0.63	1.29

Table 3. Convergence factor (ϱ) for the two–level and multilevel SA spectral AMGe method with polynomial smoother at the finest level ($\nu = 6$) and block Gauss–Seidel at coarser levels with blocks corresponding to aggregates. Fixed jump of 10^{12} in the PDE coefficient and variable fine-grid mesh. The last column shows the operator complexity of the multilevel method; namely, the sum of the non-zero entries of all level matrices divided by the number of non–zero entries of the finest-level matrix.

fixed fine-grid mesh size. Table 2 compares the behavior of the multilevel (ML) method versus the respective TL one. As we can see, the multilevel convergence factor seems also quite insensitive with respect to the contrast. Finally, Table 3 illustrates the performance of the TL and ML methods when we vary the fine–grid size and keep the contrast fixed. The TL convergence factors tend to stabilize when refining the mesh, whereas the ML ones exhibit some growth.

In conclusion, all TL-results are in good agreement with our theory; namely, the TL convergence factor stays bounded with both the contrast and is mesh-independent. Note that, all test were performed as stationary iterative method (not in a preconditioned CG iteration). If the method is used as a preconditioner in CG the convergence will be much better.

References

[ABHT] M. Adams, M. Brezina, J. Hu, and R. Tuminaro, "*Parallel multigrid smoothing: polynomial versus Gauss–Seidel,*" Journal of Computational Physics **188**(2003), pp. 593–610.

[BHMV] M. Brezina, C. Heberton, J. Mandel, and P. Vaněk, "*An iterative method with convergence rate chosen a priori,*" UCD CCM Report **140**(1999), available at http://www-math.cudenver.edu/ jmandel/papers/.

[BVV] M. Brezina, P. Vaněk, and P. S. Vassilevski, "*An improved convergence analysis of smoothed aggregation algebraic multigrid,*" Numerical Linear Algebra with Applications (2011) (to appear).

[ρAMGi] T. Chartier, R. Falgout, V. E. Henson, J. Jones, T. Manteuffel, S. McCormick, J. Ruge, and P. S. Vassilevski, "*Spectral AMGe (ρAMGe),*" SIAM J. Scientific Computing, 25(2003), pp. 1–26.

[ρAMGii] T. Chartier, R. Falgout, V. E. Henson, J. Jones, T. Manteuffel, S. McCormick, J. Ruge, and P. S. Vassilevski, "*Spectral element agglomerate AMGe,*" in Domain Decomposition Methods in Science and Engineering XVI, Lecture Notes in Computational Science and Engineering, Springer-Verlag, Berlin-Heidelberg, **55**(2007), pp. 515–524.

[GE10] J. Galvis and Y. Efendiev, "*Domain decomposition preconditioners for multiscale flows in high-contrast media,*" Multiscale Modeling and Simulation **8**(2010), pp. 1461–1483.

[LV08] I. Lashuk and P.S. Vassilevski, "*On some versions of the element agglomeration AMGe method,*' Numerical Linear Algebra with Applications **15**(2008), pp. 595–620.

[SVZ] R. Scheichl, P. S. Vassilevski, and L. T. Zikatanov, "*Weak approximation properties of elliptic projections with functional constraints,*" Available as LLNL Report LLNL-JRNL-468811, February 3, 2011.

[VBT] P. Vaněk, M. Brezina, and R. Tezaur, *"Two-grid method for elasticity on unstructured meshes,"* SIAM Journal on Scientific Computing **21**(1999), no. 3, pp. 900–923.

[Va08] P. S. Vassilevski, "*Multilevel Block Factorization Preconditioners,* Matrix-based Analysis and Algorithms for Solving Finite Element Equations," Springer, New York, 2008. 514 p.